

# About this Data Set

---

The *AIDS Public Information Data Set* is computer software designed to run on an IBM PC or compatible microcomputer, and contains information abstracted from acquired immunodeficiency syndrome (AIDS) cases reported in the United States. The data set is created each year by the Division of HIV/AIDS Prevention, National Center for HIV, STD, and TB Prevention, Centers for Disease Control and Prevention(CDC) and contains information extracted from CDC's national AIDS surveillance data base.

The December 1995 edition changes both the contents of the data set and the operation of the software. Additional data are now available for all metropolitan areas with 500,000 or more population, for metropolitan areas with 100,000 or more population from selected states, and for individual counties or health districts from selected states. Additional information is also available on patient survival time. To further protect the confidentiality of the data, some information has been removed from the data set. Month of death, vital status, central versus outlying portion of metropolitan areas, reporting delay adjustments for death dates, and information on individual AIDS-indicator conditions have been removed from the data set.

Software operation has also changed, to be more compatible with Microsoft Windows conventions and to simplify use of the software. The list of states and metropolitan areas has been modified to include counties/health districts, search criteria used to generate reports can now be saved and reused later, reports can now contain subtotals, cursor movement and item selection have been simplified, certain user prompts can be turned off, and many of the screen graphics and color combinations have been modified.

The *AIDS Public Information Data Set* contains data in two formats. The first format consists of a rectangular data file of 16 variables extracted from CDC's national AIDS data base. One-way and two-way cross tabulations of any of these variables can be displayed on your computer screen. The second format consists of a series of state, metropolitan statistical area (MSA), and county/health district tables, containing information on 8 variables included in the rectangular data file plus a location variable. There is one set of tables for the entire United States, one set for each state, one set for each MSA, and one set for each county/health district. The rectangular data file, without the state or MSA tables, is also available as an ASCII data file.

To request a copy of this data set, contact the Statistics and Data Management Branch, Division of HIV/AIDS Prevention, Mailstop E-48, Centers for Disease Control and Prevention, Atlanta, GA, 30333, telephone (404) 639-2020. Copies are also available from the CDC National AIDS Clearinghouse, telephone (800) 458-5231. If you are calling the Clearinghouse, please ask for inventory number D206. You can also download the software from the internet by linking to [http://www.cdc.gov/nchstp/hiv\\_aids/software/apids/apidssof.htm](http://www.cdc.gov/nchstp/hiv_aids/software/apids/apidssof.htm).

This manual describes the data set. It is divided into three sections and 3 appendices. On-line help screens provide additional information.

Section 1, *AIDS Surveillance in the United States*, describes the data collection process and the effect changes in this process may have on data analysis and interpretation. The section reviews the source of AIDS surveillance data and describes which patients are included in the CDC definition. It also discusses reporting delays and reporting completeness.

Section 2, *Data File Variables and Coding Schemes*, lists the variables included on the rectangular data file and describes each variable's coding scheme.

Section 3, *State, MSA, and County Tables*, describes the variables included on the state, MSA, and county/health district tables.

*Appendix A: Installation*, describes how to load and run this program on your computer. It also suggests computer hardware and software you can use to analyze the data.

*Appendix B: Metropolitan Statistical Areas* lists the MSAs included in the data set.

*Appendix C: Health Districts* lists the counties which comprise each health district included in the data set.